(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 061 487 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
20.12.2000 Bulletin 2000/51

(51) Int Cl.7: G08B 13/194

(21) Application number: 99830376.2

(22) Date of filing: 17.06.1999

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(71) Applicant: ISTITUTO TRENTINO DI CULTURA
I-38100 Trento (IT)

(72) Inventors:
- Aste, Marco
38068 Rovereto (Trento) (IT)
- Boninsegna, Massimo
38037 Predazzo (Trento) (IT)

(74) Representative: Bosotti, Luciano et al
c/o JACOBACCI & PERANI S.p.A.
Corso Regio Parco, 27
10152 Torino (IT)

(54) A method and device for automatically controlling a region in space

(57) The monitored region (S) is monitored by image signal generating means such as video cameras (2) in order to obtain a succession of images of the bodies (A, B) present in the monitored region, each image corresponding to a defined instant. The images are processed in such a way as to obtain, for each instant considered, a volumetric map of each body present in the region (S). This map, which identifies characteristics of shape, position, volume and dimensions of the body to which it refers, is processed in order to extract from it at least one parameter selected from the following group: descriptors of shape and volume, such as the volumetric map itself, the co-ordinates of position and the dimensions of each body to which the volumetric map refers. The parameter or a succession of values of the parameter obtained in this way is then compared with at least one model of these characteristics stored in a processing unit (1). Depending on the outcome of this comparison operation, a procedure of surveillance and/or reporting may be selectively activated. The solution is applicable, for example, to the automatic monitoring of museum environments, e.g. to ensure that visitors do not come too close to an exhibited work, or to the monitoring of industrial environments, e.g. to ensure that an operator does not come too close to a dangerous machine or process.

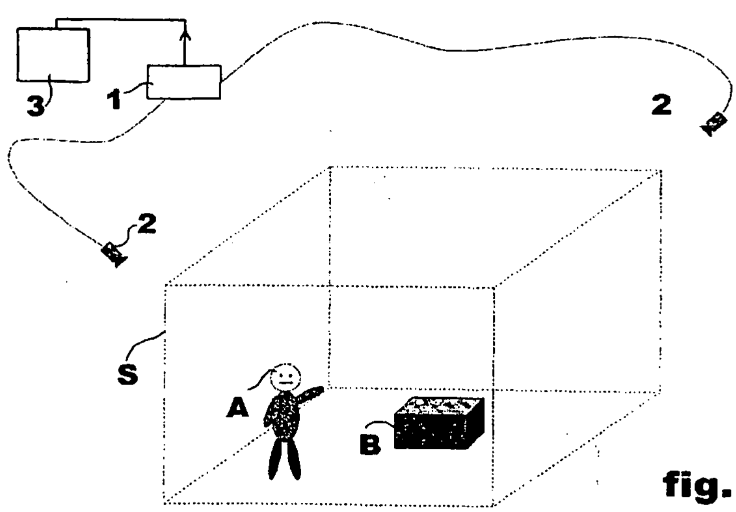


fig.1

EP 1 061 487 A1

able by the method in order to define different degrees of danger and/or alarm as a consequence of the presence of bodies within defined subregions of space.

[0014] The invention will now be described, purely by way of non-restrictive example, with reference to the attached drawings, in which:

- Figure 1 shows diagrammatically the characteristics of the system according to the invention used for the monitoring and automatic surveillance of a defined region of space,
- Figure 2, comprising four parts respectively labelled a1-a2 and b1-b2, illustrates the generation of image signals within the context of the solution illustrated in Figure 1,
- Figures 3 and 4 illustrate, in ways basically identical to those of Figures 1 and 2, another possible embodiment of the solution according to the invention; in particular, Figure 4 is composed of eight parts respectively labelled a1-a2, b1-b2, c1-c2 and d1-d2,
- Figure 5 illustrates the methods adopted for calculating a so-called map of volumetric occupation,
- Figure 6 illustrates schematically one of these maps capable of being obtained within the context of the invention, and
- Figure 7 is a flow diagram relating to the generation and use of such a map.

[0015] In particular the expression "volumetric map" as used here means any representation of occupied volumes due to the presence of a body, in other words a representation of a three-dimensional map in which the regions of volumetric occupation introduced by the presence of bodies are indicated. Such a map is obtained after image analysis procedures have been carried out using automatic methods known per se or according to the embodiments of the invention described below. For a summary of some of these methods the following may usefully be referred to: Marr D., "Vision", Freeman, 1982; Ballard D.H. and Brown C.M. "Computer Vision", Prentice Hall, 1982; Martin W.N. and Aggarwaal J.K., "Volumetric description of objects from multiple views", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 5, pp.150-158, 1983. From this map, by means of the volumetric analysis carried out using automatic methods known per se it is possible to derive the characteristics of shape, volume, dimensions and position of the bodies present in a defined region of space in such a way that they can easily be compared with similar representations obtained from the volumetric maps of other bodies.

[0016] In both Figure 1 and Figure 3, the reference S indicates a region of space in which it is wished to detect the presence of people A or objects B.

[0017] The region S may be bounded by physical walls , as for example in the case of a room or cage, or may consist simply of a portion of space bounded by an imaginary closed surface that separates a generic space in two regions, or it may be bounded partly by physical barriers, for example the floor, and partly by an imaginary surface. The monitored region has however the feature of a volume, may be of any shape and can be defined simply and flexibly according to need.

[0018] In the currently preferred embodiment of the invention, the volumes occupied by the bodies (such as bodies A, B visible in Figures 1 and 3) that are present in the monitored region S are found by using two or more video cameras (acting as image signal generating means) installed in such a way that the region S is in the visual field of at least two video cameras 2, as shown for example in Figures 1 and 3 (the latter figure referring to a solution in which four video cameras 2 are used). It is advisable for the video cameras 2 to be so positioned as to avoid occlusions due to the movement of objects on the same plane; for example, for bodies of different heights standing on the floor and moving about, it is preferable to have views from above.

[0019] The signals (of analogue type or already directly converted into digital form) output by the video cameras 2 are sent to a processing unit 1 which may be a specialized processor or, in the currently preferred embodiment of the invention, a computer such as a programmed personal computer (known per se) in order to extract from the images the shapes of the bodies A and B present within the region S to be monitored. The object here is to check for the possible presence of bodies not inherently belonging to the monitored region.

[0020] In particular, Figures al and bl included in Figure 2, and Figures al, bl, cl and dl included in Figure 4 show the images produced by the two video cameras depicted in Figure 1, on the one hand, and by the four video cameras depicted in Figure 3, on the other.

[0021] Using the signals corresponding to the abovementioned images, the unit 1 is able, in accordance with known principles (typically by using known image processing algorithms), to extract respective sets of data representing the abovementioned shapes of the bodies present within the region S. For example, Figures a2 and b2 of Figure 2, and Figures a2, b2, c2 and d2 of Figure 3 show the shapes, marked C of the body A present within the region S corresponding to images a1 and b1, or a1, b1, c1 and d1, produced, from their respective points of observation, by the video cameras 2.

[0022] For an overview of these algorithms, the following may usefully be referred to: Huang T.S., "Image sequence processing and dynamic scene analysis", Springer-Verlag, 1982; Jain A., "Fundamentals of digital image processing", Prentice Hall, 1989; Jain J.R. and Martin W.N. and Aggarwaal J.K., "Segmentation through the detection of changes due to motion", Computer Graphics and Image Processing, vol. 11, pp. 13-34, 1979; Debuisson M.-P., "Contour extraction of moving objects in complex outdoor scenes", Int. Journal of Computer Vision, vol. 14, pp.83-105, 1995; Bichsel M., "Segmenting Simply Connected Moving Objects in a Static Scene", IEEE Transactions on Pattern Analysis

which the monitored space has been divided are occupied by objects. The aim of all this is to obtain, as a result, the representation seen in Figure 6. Using this technique, the volumetric map approximates to the volume of actual occupation with a resolution based on the dimensions of the cells D1, D2 and on the spatial resolution of the means employed to generate the image signals (in particular, in the case of video cameras composed of a matrix of sensitive points, the resolution in a particular direction is given by the ratio of the dimension of the observed region to the number of sensitive elements in that direction). The dimensions of the cells are defined having regard not only to the resolution but also to the processing capacity of the unit 1 and of the frequency with which the surveillance information is to be updated. Also borne in mind is the fact that the overall degree of approximation can be enhanced, as already stated, by using more video cameras in suitable positions.

[0035] In particular, the processing unit 1 can be programmed, again in a known manner, to carry out a volumetric analysis of bodies for the purpose of recognizing the distinctive characteristics of shape and volume of the objects analysed. The programming can be done by conventional algorithmic approaches, thus coding ab initio the characteristics of shape and volume of the bodies to be detected into the processing system, or using statistical learning techniques such as for example neural networks. It is also possible to design the unit 1 such that it is able to evaluate the way the positions of the bodies examined within the region S are changing in space and time and deduce the dynamics of their movements, in particular the line, and the direction along this line, of their displacements.

[0036] This point will become clearer on referring to the flow diagram given in Figure 7, which shows, in a deliberately schematic way, for ease of comprehension, the principles by which the functions of automatic monitoring are carried out in the unit 1.

[0037] Assuming the process to start at a starting step 100, in a step 101 the unit 1 examines the data set corresponding to the images generated by the video cameras 2 (optionally already processed to refer only to moving objects) and in a second step 102 commences an action of scanning the region S in such a way as to scan the cells D into which the region S has been theoretically divided up. As a general rule, each of these cells will be identified by three co-ordinates x_i, y_i, z_i within the system x, y and z to which the bottom part of Figure 5 refers.

[0038] From now on it will be assumed, for simplicity, that this scanning operation applies, on each successive detection of the images of the region S, to all the cells contained within the region S scanned on a "matrix" principle, for example in successive lines (co-ordinate x), columns (co-ordinate y) and planes (co-ordinate z).

[0039] Those skilled in the art of image processing will have realized that it is possible (e.g. in order to reduce the processing cost and/or speed up the processing) to adopt different scanning systems, such as predictive-type scanning systems which, once initialized with reference to a map of initial volumetric occupation, perform subsequent scans only on cells where there exists some degree of likelihood inherent in the fact that these cells may be significant in the generation of subsequent maps, the aim being to avoid the need to perform exhaustive scanning of the entire region S for each updating operation.

[0040] In this context it is also known that it is possible to intervene in such a way that, when operating on the abovementioned principles, the unit 1 is also capable of detecting, for example, the entry into the region S of a body not previously present, the aim being to extend the scanning action to those cells (previously not included in the scanning action) which the body subsequently occupies.

[0041] The steps marked 1031, 1041; 1032, 1042; ...; 103n, 104n indicate successive processing stages, here shown as carried out in parallel, though in fact they can be performed serially, and therefore sequentially in time. In the course of these steps, for each video camera 1, ..., n (n is equal to 2, and to 4, in the illustrative embodiments shown in Figures 1 and 3, respectively) and for each cell $D(x_i, y_i, z_i)$ that is scanned, the unit 1 checks to see whether the cells corresponding to their respective images generated by the video cameras 2 can be regarded as occupied or unoccupied.

[0042] In the next step, indicated by the general reference 105, the results of the comparisons carried out in steps 1041, 1042, ..., 104n are processed in order to decide whether, on the basis of the image data, the scanned cell is to be regarded as occupied or unoccupied for the purposes of constructing the map of volumetric occupation.

[0043] The relevant criteria for attributing the "occupied" or "unoccupied" logic value may differ.

[0044] On this subject it should be remembered that the cells of the region S are not necessarily all covered by all of the video cameras 2. As a consequence, in the case of certain cells, attribution of the "occupied" value may be based on a different number of decision processes relating to the individual images than the number of images taken into consideration in attributing the "occupied" logic value to other cells.

[0045] The criterion used in attributing the logic value in question may be of unanimous type (the cell is judged to be occupied for the purposes of the construction of the map of volumetric occupation if and only if all the video cameras 2 whose images are taken into account produce data corresponding to occupation in the relevant image), majority type (the cell is judged to be occupied if the majority of video cameras 2 give data indicating occupation in the respective images), or correlation with the values attributed to adjacent cells (so that uncertainty in the attribution of the "occupied" value to a cell is resolved on the basis of confident values attrib-

be people only.

[0057] Furthermore, it is possible to detect the simultaneous presence of several bodies, even if of different kinds, in the monitored region. The manner in which the position of the bodies change within the monitored region can be used to detect violation of predefined subregions. It is thus possible to monitor, as has already been seen, the presence of a movement of people in the vicinity of a machine in an industrial environment and activate an alarm signalling procedure whenever at least one person comes within a certain distance of that machine.

[0058] The solution described is highly robust and overcomes the functional limitations of currently used systems. Thus, it is capable of detecting the presence and at the same time determining the position of people or objects within a defined region of space, discriminate between objects and people, between objects or people close to each other, and between objects and people that move into the monitored region following different paths or more generally with behaviours which could easily deceive other types of sensor.

[0059] Those skilled in the art will recognize that the method according to the invention can be carried out using, at least in part, a computer program capable of being run on a computer in such a way that the system comprising the program and the computer carries out the method according to the invention. The invention therefore extends also to such a program capable of being loaded into a computer which has the means of or is capable of carrying out the method according to the invention, as well as to the corresponding information technology product comprising a means readable by a computer containing codes for a computer program which, when the program is loaded into the computer, cause the computer to carry out the method according to the invention.

[0060] Clearly, without affecting the principle of the invention, the constructional details and the embodiments may be greatly altered compared to what has been described and illustrated, without thereby departing from the scope of the present invention, as defined in the accompanying claims.

Claims

1. Method for the detection and location of bodies (A, B) in a defined region of space (S), comprising the operations of generating (2) image signals capable of representing a succession of images of at least one body present in the said region (S), each image corresponding to a defined instant, characterized in that it comprises the following operations:

   - processing (101 to 106) the said image signals in such a way as to obtain for each instant taken into consideration a volumetric map (F) of the said at least one body present in the said region (S), the said volumetric map (F) representing the shape, position, volume and dimensions of the body to which the said volumetric map (F) refers,
   - extracting (107) from the said volumetric map (F) at least one parameter (P) taken from the following group: descriptors of shape and volume, such as the volumetric map (F) itself, the position co-ordinates and the dimensions of the said at least one body to which the said volumetric map (F) refers,
   - comparing (108) the said at least one parameter (P) with at least one model (D) for compatibility of the said volumetric map (F) with predetermined conditions of occupation of the said region (S), and
   - selectively generating a warning signal (109) depending on the outcome of the said comparison (108).

2. Method according to Claim 1, characterized in that it comprises the operations of storing volumetric maps (F) or successions of the said at least one parameter (P) obtained from image signals relating to images of the said succession corresponding to successive instants, in order to detect changes in time in the said volumetric map (F) or the said at least one parameter (P), and in that the said model is itself generated as a model of changes over time.

3. Method according to Claim 1 or Claim 2, characterized in that it comprises the operation of comparing (108) successions of the at least one parameter (P) obtained from image signals relating to images of the said succession corresponding to successive instants with at least one model for compatibility of said successions with predetermined conditions of occupation of the said region (S).

4. Method according to Claim 1 or Claim 2, characterized in that it comprises the following operations:

   - generating image signals relating to the view of the said region (S) from separate observation points (2) so as to generate at least two separate image signals relating respectively to the projections of the same points of the said region (S) viewed from separate observation points,
   - processing (1) the said separate image signals by finding the match between the projections of the same real-world points onto separate images with a view to finding its position in space,
   - obtaining the said volumetric map (F) from the said positions in space.

5. Method according to Claim 1 or Claim 2, characterized in that it comprises the following operations:

Claims 1 to 5.

13. Information technology product comprising a means readable by a computer containing codes for a computer program which, when the program is loaded into the computer, cause the computer to carry out the method according to any one of Claims 1 to 5.

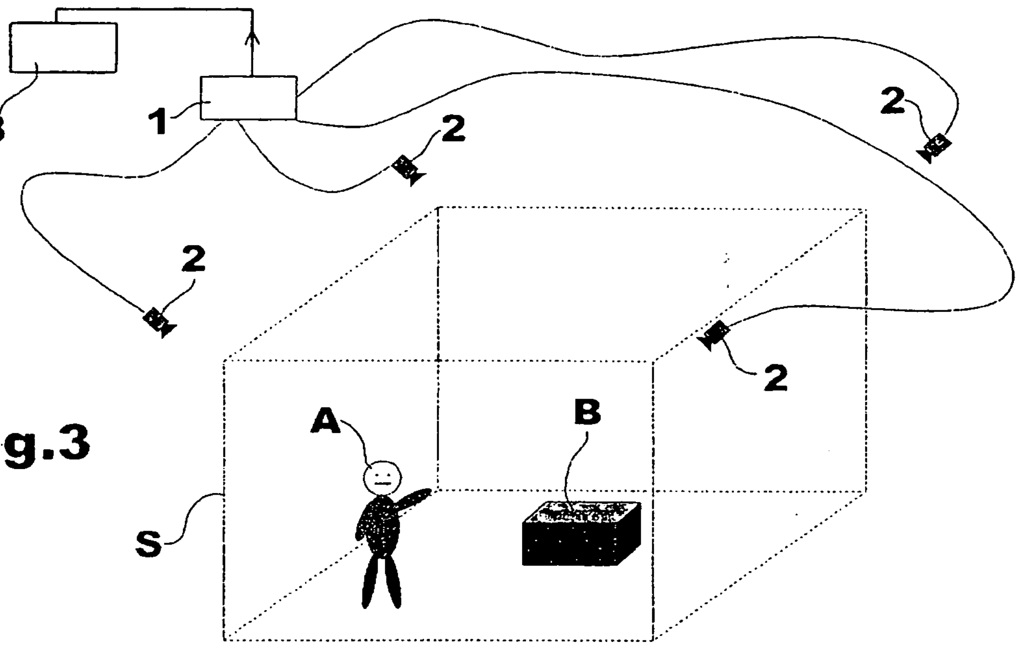


fig.3

fig.4

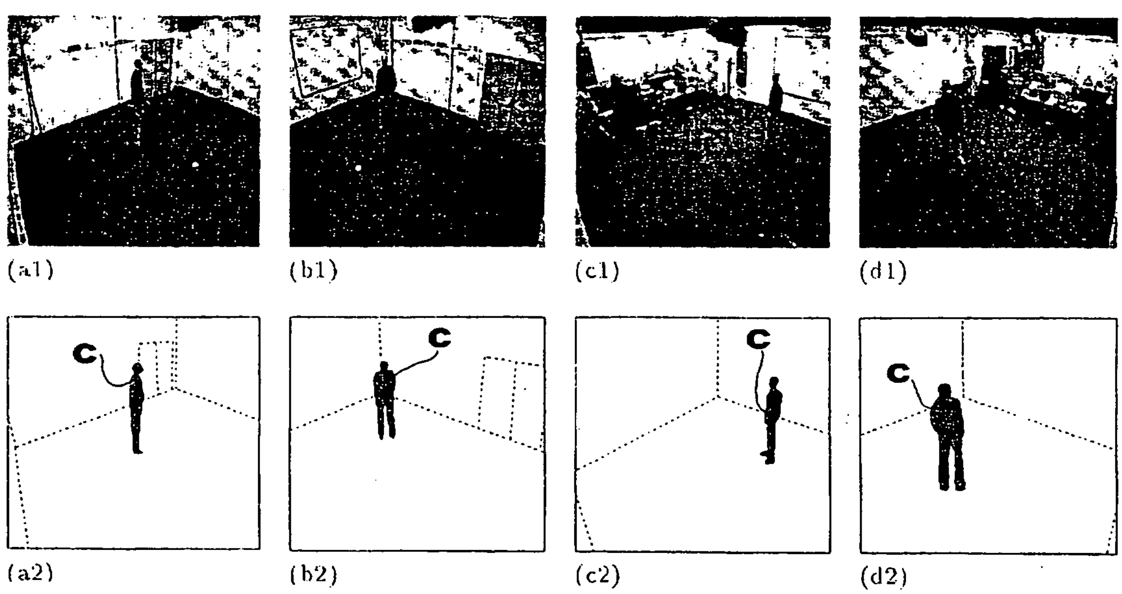
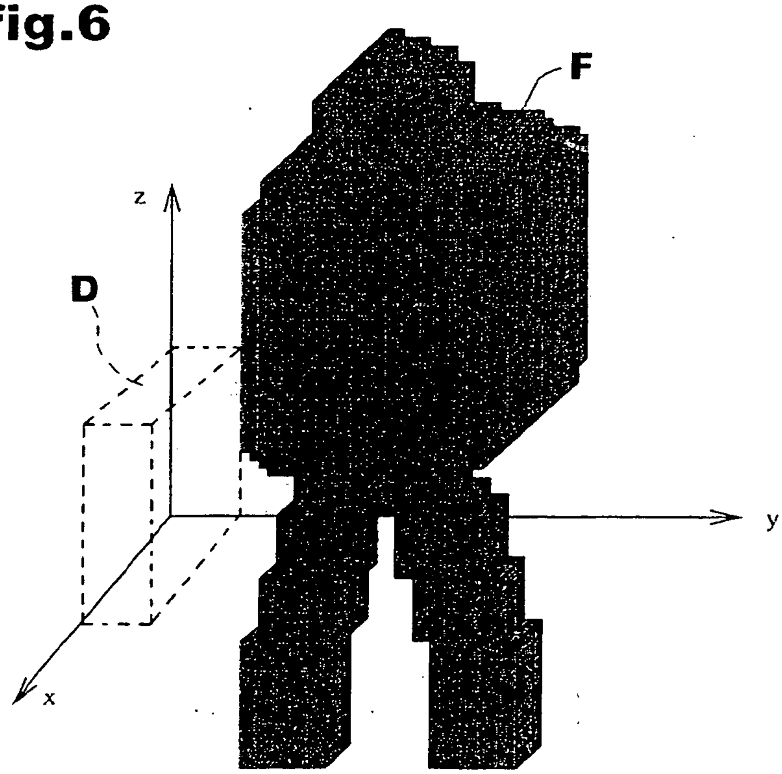


fig.6



European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 99 83 0376

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
Y	EP 0 356 734 A (SIEMENS AKTIENGESELLSCHAFT) 7 March 1990 (1990-03-07) * column 2, line 53 - column 3, line 5 * ---	1,6	G08B13/194
Y	TARBOX G H ET AL: "VOLUMETRIC BASED INSPECTION*" PROCEEDINGS OF THE IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS,US,NEW YORK, IEEE, vol. -, page 1239-1246 XP000334081 ISBN: 0-7803-0738-0 * abstract * ---	1,6	
A	CARLSON J ET AL: "Real-time 3D visualization of volumetric video motion sensor data" SURVEILLANCE AND ASSESSMENT TECHNOLOGIES FOR LAW ENFORCEMENT, BOSTON, MA, USA, 19-20 NOV. 1996, vol. 2935, pages 69-79, XP000863382 Proceedings of the SPIE - The International Society for Optical Engineering, 1997, SPIE-Int. Soc. Opt. Eng, USA ISSN: 0277-786X * abstract * -----	1,6	

TECHNICAL FIELDS SEARCHED (Int.Cl.7)	G08B

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	16 December 1999	Chateau, J-P

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)